

**SYSTEMS AND METHODS FOR ROBUST, REAL-TIME MEASUREMENT
OF NETWORK PERFORMANCE**

Inventors:

Michael Lloyd

Sean Finn

Omar Baldonado

Mansour Karam

Faisal Siddiqi

James McGuire

Herbert Madan

Prepared by Wilson Sonsini Goodrich Rosati

0903423, 07/001

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Applications 60/241,450, filed October 17, 2000 and 60/275,206, filed March 12, 2001, which are hereby incorporated by reference in their entirety.

5

Description of Related Art

The performance characteristics of routes in internetworks, such as the Internet, have been assessed in prior efforts. Statistical metrics of Internet performance include the characteristics of jitter, loss, and delay. Jitter may be characterized as the amount of variance in the time taken by packets traversing a path in a network. Delay indicates the amount of time taken for packets to traverse the path. And loss indicates the lossiness of the internetwork path.

Empirical observations have demonstrated that various combinations of these performance metrics are especially relevant to the performance of certain types of applications on the Internet. For instance, in some voice streaming applications such as Voice over IP (VoIP), appreciable levels of jitter may have a highly deleterious effect on performance, while some packet loss may be tolerable. In other applications, jitter and delay may be tolerable, while significant packet loss may be fatal.

Given the significance of such metrics to Internet performance, there is a need to measure such statistics in real-time for arbitrary end-points in an internetwork. The prior art also evinces a need to ensure that such statistics are robust, and based on substantial packet traffic.

Summary of the Invention

Some embodiments of the invention include methods and apparatuses for obtaining delay, jitter, and loss statistics of a path between server and an end user coupled via an internetwork; in some embodiments, the server may comprise a web server in communication with the end user via the Internet. In some embodiments of the invention, these statistics are obtained by analyzing the details of a TCP connection underlying an HTML transaction. Some such embodiments ensure robust measurements of jitter, delay, and loss by maximizing traffic between the web server and the surfer in order to generate a robust sample of TCP connections.

In some such embodiments, content is updated with one or more html link(s). This existing content may reside on a highly trafficked portal, such as a web portal, and may be encoded in a markup language, such as Hyper Text Markup Language (HTML). The Uniform Resource Locators (URLs) corresponding to the one or more links resolve to the server from which the statistics are to be measured, i.e., the server which connects to the end user over the desired path. In some embodiments, this resolution may be based on an explicit relationship between a URL and a given measurement path. In alternative embodiments, the one or more URLs may resolve to an address which varies on each invocation, such that only the address, rather than the URL, connotes a relationship with the specific measurement path. A request for the connection comes into the server, and based on the target address, the outbound response is subsequently forced to a specific measurement path. In some embodiments of the invention, the actual content supplied by the server is minimized, in order to preserve bandwidth. In some embodiments, the content may be visually

imperceptible, comprising one or more pixels, which may be transparent. In other embodiments, the content may comprise a visual artifact

Some embodiments of the invention include a measurement subsystem which records observed call response times, which are used to record round trip times for packets traversing the path between the server and the end user. In some embodiments, these packets employ the TCP/IP protocol for their transport. In alternative embodiments, these measurements may be gathered at the end-user side, as opposed to the server side.

Some embodiments of the invention measure round trip times for different patterns of TCP messages sent within a TCP connection. In some embodiments, these measurements of round trip times are converted into measurements of jitter, loss, or delay along the desired path. In some embodiments of the invention, jitter, loss, and delay statistics may be inferred by groups, or classes, of end user addresses. These and other embodiments are further described herein.

Brief Description of Figures

Figure 1 illustrates an architecture used to redirect internetwork traffic to a measurement server according to some embodiments of the invention.

5

Figure 2 illustrates techniques used to measure Round Trip Times for various types of TCP sessions according to some embodiments of the invention.

0903423 071001

Detailed Description

Distributing Hits to the Desired Server

Some embodiments of the invention include systems and methods to maximize traffic through a desired path, in order to generate a robust number of measurements of round trip times through the path. These embodiments are illustrated schematically in Figure 1. The method generates traffic towards an end user 102, or surfer. An internetwork 100 includes a measured server 104, which is the server from which traffic is to be measured, and a highly trafficked portal 106. The highly trafficked portal 106 may include content from a popular commercial web site. The measured server and the end user can communicate via the internetwork through one or more paths 108. Such embodiments attempt to divert traffic from the portal 106 to the measured server 104, in order to ensure robust measurements of network performance along the one or more paths 108.

In some such embodiments, a content object is included in the portal 106, so that when an end user 102 connects to the portal 106, her request is redirected to the measured server 104 in order to receive the portion of content. This content object may be referred to as a webby. In some embodiments of the invention, the webby is designed to occupy a minimal amount of bandwidth. In some embodiments, the webby is designed to be imperceptible. In a non-limiting implementation of the webby, the content object may comprise a transparent GIF or JPEG, which includes one or more pixels. Other implementations of the content object will be apparent to those skilled in the art.

In web based embodiments, when a surfer's browser 102 requests the content object, the browser 102 performs a DNS lookup, and retrieves an IP address for the web object; this IP address resolves to the measured server 104.

In some embodiments of the invention, by supplying varying answers for the IP address, hits may be distributed across many measured servers 104. In response to the request, the measured server 104 delivers the content object to the surfer's browser 102.

5

Measuring Round Trip Times

Some embodiments of the invention measure Round Trip Times (RTTs) between the measured server 104 and end users 102 in order to generate metrics of path performance; these metrics may, by way of non-limiting example, include jitter, delay, and loss statistics. In some embodiments of the invention, different algorithms for measuring RTTs are employed, contingent upon the type of session that is witnessed. As such, several types of TCP sessions are described herein, followed by a discussion of the RTT measurement techniques that may be employed for the various sessions. Note that the discussion that follows employs acronyms described in Table 1 below:

Table 1 Acronyms used in the description of TCP patterns

Si	SYN received by the webby (i.e., incoming SYN)
So	SYN/ACK sent by the webby
Pi	PUSH packet received by the webby
Po	PUSH packet sent by the webby
Fi	FIN message received by the webby
Fo	FIN message sent by the webby
.i	ACK message received by the webby
.o	ACK message sent by the webby

Figure 2 illustrates three types of sessions 200 202 204 that may be witnessed between the measured server 104 and the end user, or surfer 102. These patterns are hereafter referred to as Basic Pattern 1 (B1) 200, Basic Pattern 2 (B2) 202, and Basic Pattern 3 (B3) 204. The differences between patterns B1 on one hand, B2 and B3 on the other, inheres in the manner in which TCP behaves on the side of the webby, i.e., the measured server 104. In the case of B1 200, the actions performed by the webby 104 upon the receipt of a PUSH packet (i.e., P_i) are as follows:

- The webby 104 sends an ACK packet acknowledging the PUSH.
- The webby 104 sends the requested data in a PUSH packet.
- The webby 104 subsequently terminates the connection by sending a FIN message.

For cases B2 202 and B3 204, the actions performed by the webby 104 upon the receipt of a PUSH packet (i.e., P_i) are as follows:

- The webby 104 sends an ACK packet acknowledging the PUSH
- The webby 104 sends the requested data in a PUSH packet
- The webby 104 subsequently waits for an acknowledgment from the surfer 102 containing notification of receipt of the data before the webby 104 proceeds with sending a FIN.

In some embodiments of the invention, Round Trip Times RTT_1 , RTT_2 and RTT_3 are computed by use of the same algorithm in all cases 200 202 204. In some such embodiments, RTT_1 may be determined simply by waiting for an ACK corresponding to the *first* SYN/ACK. In some embodiments, RTT_2 may be measured by starting a timer at the instant the first PUSH is sent by the webby 104 (as for RTT_1 , the timer is started at the *first* PUSH to take into account the

effect of timeouts), and stopping the timer upon the receipt of the first packet acknowledging the PUSH that was sent. (This packet acknowledges a sequence number at least equal to that of the PUSH message). A similar technique may be applied to RTT_3 , this time to the FIN packet sent by the webby 104. As discussed in U.S. Provisional Applications 60/241,450, filed October 17, 2000 and 60/275,206, filed March 12, 2001, which are hereby incorporated by reference in their entirety, these techniques for measuring Round Trip Times have been empirically shown to be robust in all manner of complex TCP transactions.

Computation of Jitter, Loss, and Delay from Round Trip Times

In some embodiments of the invention, a measurements listener receives values of RTT_1 , RTT_2 , and RTT_3 that correspond to a given IP address. In some embodiments, the measurements listener may comprise one or more processes distributed on one or more servers coupled to the internetwork. These measurements are sent to a module that performs one or more of the following steps:

- **Compute the values of round-trip time d , jitter v , and packet loss p for this measurement instance.**
- **Map the IP address to a corresponding group of IP addresses** (this group may comprise an Equivalence Class, which is further described in which are hereby incorporated by reference in their entirety)
- **Update the values of \hat{d} , \hat{v} , \hat{p} , using old values of \hat{d} , \hat{v} , \hat{p} and the values of d , v , and p , wherein \hat{d} , \hat{v} , \hat{p} comprise weighted averages of delay, jitter, and loss, respectively.**

Non-limiting implementations for calculating d , v , and p from the Round Trip Times are described herein. First, note that RTT_1 and RTT_3 do not overlap in some embodiments. Hence, network events that are captured by the first round trip time RTT_1 are typically not captured by RTT_3 . Empirical observations also demonstrate that RTT_1 and RTT_3 are often very different. As such, some embodiments of the invention employ a difference between RTT_1 and RTT_3 to capture network oscillations in performance, i.e. jitter. In one such embodiment the jitter, v is set to the absolute value of the difference, i.e.,

$$v = |RTT_3 - RTT_1|$$

Empirical observations also demonstrate that RTT_2 and RTT_3 may be highly correlated. As such, in some embodiments of the invention a difference between RTT_2 and RTT_3 may be used to infer packet loss. In case RTT_3 is not measured, a large difference between RTT_1 and RTT_2 may be used to infer packet loss in extreme cases, for example when RTT_1 is close to 0, and RTT_2 has a value on or about 3 seconds. Otherwise, a difference between RTT_2 and RTT_3 that is close to 3 or 6 seconds may be used in some embodiments of the invention, to declare packet loss. Thus, to determine loss, some embodiments of the invention employ one or more of the following steps:

- If either RTT_1 or RTT_2 is small (for example, less than 500 ms), compute the difference between RTT_1 and RTT_2 : if this difference is on or about 3 seconds or 6 seconds, set p to 1.
- If either RTT_1 or RTT_2 is large (for example, more than 500 ms), compute the difference between RTT_2 and RTT_3 : if this difference is on or about 3 seconds or 6 seconds, set p to 1.

- Otherwise set p to 0.

In some embodiments of the invention, d is set to an average of the *true* RTTs measured for a transaction. In case p is set to 0, this is simply the average of all three RTTs. In case p is set to 1, the packet involved in the loss should be removed from the computation of the average d . (Alternatively, a 3 second timeout can be subtracted from the measured *RTT* for that packet.)

As will be apparent to those skilled in the art, the implementations described are non-limiting techniques for computing d , v , and p from Round Trip Times; other implementations will be apparent to those skilled in the art.

Computing Weighted Averages of Jitter, Delay, and Loss

Some embodiments of the invention include techniques for maintaining weighted averages of Delay, Jitter, and Loss, \hat{d} , \hat{v} , and \hat{p} respectively. In some such embodiments, current values of d , v , and p values as well as previous values of \hat{d} , \hat{v} , and \hat{p} for a relevant group of IP addresses are used to compute the new values for \hat{d} , \hat{v} , and \hat{p} .

In a non-limiting example, weighted moving averages are used to compute \hat{d} , \hat{v} , and \hat{p}

$$\begin{aligned}\hat{d}_{new} &= \alpha \hat{d}_{old} + (1 - \alpha)d \\ \hat{v}_{new} &= \beta \hat{v}_{old} + (1 - \beta)v \\ \hat{p}_{new} &= \gamma \hat{p}_{old} + (1 - \gamma)p\end{aligned}$$

In some embodiments, α , β , and γ are fixed constants. In some such embodiments, the combination of values used for α , β , and γ are determined by the type of application the TCP session is supporting. These applications may include, but are not limited to, any one or more of HTTP 1.0, HTTP 1.1, Voice over IP, or Video streaming over IP. Examples of values of α , β , and γ that may be used for these applications are presented below in an XML format. Note that these examples also include sample values for parameters denoted theta, phi, omega, and psi; these parameters may be used to convert the tuples (α , β , and γ) into a scalar performance score; these parameters are further described in U.S. Provisional Applications 60/241,450, filed October 17, 2000 and 60/275,206, filed March 12, 2001, which are hereby incorporated by reference in their entirety. The values presented herein are for illustration only; other value combinations will be apparent to those skilled in the art:

HTTP 1.0

```
<module> <engine slot="1"> <application model="http1.0" [alpha="0.9" beta="0.9"
gamma="0.9" theta="1.18" phi="0.13" omega="0.15" psi="0.25"] /> </engine>
</module>
```

HTTP 1.1

```
<module> <engine slot="1"> <application model="http1.1" [alpha="0.9"
beta="0.9" gamma="0.9" theta="1.3" phi="0.31" omega="0.41" psi="1.0"] />
</engine> </module>
```

Voice over IP

```
<module> <engine slot="1"> <application model="voice" [alpha="0.9" beta="0.9"
gamma="0.9" theta ="1.5" phi="6.0" omega="23.0" psi="0.0"] /> </engine>
</module>
```

5

Video Streaming

```
<module> <engine slot="1"> <application model="video" [alpha="0.9" beta="0.9"
gamma="0.9" theta="1.0" phi="4.0" omega="69.0" psi="0.0"] /> </engine>
</module>
```

In some embodiments of the invention, time-decaying values of α , β , and γ may be employed. In some such embodiments, these values of α , β , and γ may decay exponentially, i.e.,

$$\alpha = \exp(-k_{\alpha} T)$$

$$\beta = \exp(-k_{\beta} T)$$

$$\gamma = \exp(-k_{\gamma} T)$$

Other value combinations for α , β , and γ shall be apparent to those skilled in the art.

Conclusion

The various techniques presented above for measuring Round Trip Times and determining jitter, loss, and delay values are presented for illustrative purposes only. Many equivalent techniques shall be apparent to those skilled in the art.